AD

EXPERIMENTAL COMPARISON OF MONTE-CARLO SAMPLING TECHNIQUES TO EVALUATE THE MULTIVARIATE NORMAL INTEGRAL

Elizabeth N. Abbe

STATISTICAL RESEARCH AND ANALYSIS DIVISION

DDC

OCT 21 1969

AE/RL

U. S. Army
Behavioral Science Research Laboratory

June 1969

AD [ ]

EXPERIMENTAL COMPARISON OF MONTE-CARLO SAMPLING TECHNIQUES TO EVALUATE THE MULTIVARIATE NORMAL INTEGRAL

Elizabeth N. Abbe

Richard C. Sorenson, Task Leader

STATISTICAL RESEARCH AND ANALYSIS DIVISION
Cecil D. Johnson, Chief
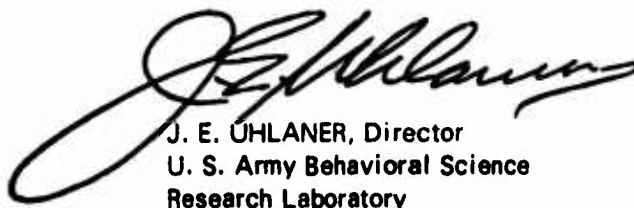
FOREWORD

BESRL's OPTIMIZATION MODELS Work Unit seeks to provide means of solution to personnel management problems relating to the distribution, training, career progression, reassignment, and utilization of personnel in current and future Army personnel subsystems. Personnel systems are analyzed and areas identified for which objective optimization techniques can profitably be applied.

Quantitative models and computing algorithms for optimal assignment and for evaluating the feasibility of alternative approaches to personnel system problems have been provided. Optimization techniques are developed to identify the "best" policy for meeting future military contingencies at minimum cost in terms of personnel system resources. The present Technical Research Note compares techniques for estimating manpower requirements where a number of individually varying skills, performance potentials, background and behavioral factors must be considered.

The entire research work unit is responsive to requirements by RDT&E Project 2Q062106A723, "Human Performance in Military Systems," FY 1969 Work Program.

J. E. UHLANER, Director
U. S. Army Behavioral Science
Research Laboratory

EXPERIMENTAL COMPARISON OF MONTE-CARLO SAMPLING TECHNIQUES TO EVALUATE THE MULTIVARIATE NORMAL INTEGRAL

BRIEF

Requirement:

To represent more realistically and precisely the complex and multiple interrelationships within the Army personnel system, continued development of quantitative and technical methodology is required. The present objective was to evaluate two different numerical methods for estimating probability when a multivariate normal model (for example, one involving scores on a battery of tests) can be assumed.

Procedure:

In a series of simulation experiments in which random vector observations were generated, probability estimates were computed by each of the methods. Probability regions on which the experiments were based were chosen to have a variety of properties. The precision of the two methods was compared from the magnitudes of the variances of the probability estimates over independent samples.

Findings:

Results appear to be affected both by the size of the probability region being estimated and by the goodness of the approximation of the sampling distribution to the unknown distribution. The more complex method was consistently superior for very small probability regions; but when the sampling approximation was poor, the precision of the probability estimates favored the simpler approach.

Utilization of Findings:

The computational procedures developed in conjunction with this series of experiments are considered practical methods of estimating probability based on multiple scores for individuals in a sample population. Estimation problems for which one method can be expected to be superior to the other were clarified. Changes in the computational procedures from which methodological improvements may be expected were also made apparent.

**EXPERIMENTAL COMPARISON OF MONTE-CARLO SAMPLING TECHNIQUES
TO EVALUATE THE MULTIVARIATE NORMAL INTEGRAL**

CONTENTS

EXPERIMENTAL COMPARISON OF MONTE-CARLO SAMPLING TECHNIQUES TO EVALUATE THE MULTIVARIATE NORMAL INTEGRAL

Personnel management research relating to manpower has led to the development of both optimization and simulation models within BESRL. These models can relate to the formulation of policies regarding man-power requirements and are effective in the evaluation of such policies. Realistic estimates require, however, that multiple behavioral response patterns, as well as the complex interrelationships among these patterns, be incorporated into the model of the personnel system. The continued development of more advanced quantitative techniques is therefore required.

The research reported is directed toward the development of one such quantitative technique for use with models for which the multiple-response patterns can be characterized by a particular, but very commonly encountered distribution, the multivariate normal distribution. One example of the application of this technique is in policy formulation, where it may be necessary to estimate the proportionate increase of men assigned in various Army job areas which would occur if requirements on certain classification criteria were made more lenient. A suggested policy of lowering requirements five points on each of the job areas under investigation might show, for instance, that too few men would be classified into these areas to meet increased manpower needs. As another example, some estimate of the proportion of the total Army population which lies between certain bounds on a system of behavioral measures might be required before performing a valid simulation experiment. Such problems require the integration of the multivariate probability dis·tribution characterizing the population, which in this case is assumed to be the multivariate normal distribution.

A general analytical method for integrating the multivariate normal distribution is at the present time unavailable. The integral can be approximated by numerical quadrature, but a very large number of data points needs to be generated when the number of dimensions is large. Another approach, designed to improve the approximation to the integral and shorten computation time, is to obtain the observations by random sampling over the region of integration rather than use the systematic sampling of quadrature. Such a "Monte-Carlo" approach will yield prob-ability estimates which vary from sample to sample. The larger the variance, however, the larger is the number of points which must be generated to obtain a given degree of precision, and again the time required for computations may be appreciable. Assuming, then, that a given Monte-Carlo method yields unbiased estimates, the precision of the method can be evaluated from the magnitude of the variance of the esti-mates over independent samples.

Experimental results have been obtained to evaluate two different methods of Monte-Carlo sampling to integrate the multivariate normal prob-ability density function. Results for different types of probability integrals and a description of the sampling techniques are presented.

MONTE-CARLO METHODS

Tallied Sampling (Monte-Carlo Method I)

One of the Monte-Carlo methods described has already been used and reported by Hillier (1). Random vector observations are generated to have the multivariate normal distribution of interest. The proportion of observations which fall within the specified region of integration constitutes the probability estimate. Even though the computations involved are simple, this type of estimate shows considerable change from sample to sample. Such variability serves, however, as a useful standard by which to compare the behavior of a second method of Monte-Carlo sampling, more complex computationally, but which has showed promise of yielding estimates with a higher degree of precision.

The Model is described below:

Since any multivariate normal distribution can be easily converted to standard form, methods of integration can be described in more general terms with reference to the standardized distribution. Let $X = [x_1, x_2, \ldots, x_k]$ represent a k-dimensional vector of random variables distributed as the multivariate normal with mean vector $\underline{0}$ and correlation matrix R. The multivariate normal integral over a given region $A = (\ell_1, u_1); (\ell_2, u_2), \ldots, (\ell_k, u_k)$ is

$$\pi = f(X)dX$$

$$= \frac{|R|^{-\frac{1}{2}}}{(2\pi)^{k/2}} \int_{\ell_1}^{u_1} \int_{\ell_2}^{u_2} \cdots \int_{\ell_k}^{u_k} \exp\left[-\tfrac{1}{2}(X\,R^{-1}\,X')\right] dX.$$

To estimate π, introduce a new random variable y, where

$$y = 1 \text{ if } X \in A$$

$$y = 0 \text{ otherwise.} \tag{1}$$

In scalar notation, $X \in A$ if $\ell_j < x_j < u_j$ for $j = 1, 2, \ldots, k$.

Then

$$E(y) = \int_{-\infty}^{\ell_1} \int_{-\infty}^{\ell_2} \cdots \int_{-\infty}^{\ell_k} y\, f(x_1, x_2, \ldots, x_k)\, dx_1\, dx_2 \ldots dx_k$$

$$+ \int_{\ell_1}^{u_1} \int_{\ell_2}^{u_2} \cdots \int_{\ell_k}^{u_k} y\, f(x_1, x_2, \ldots, x_k)\, dx_1\, dx_2 \ldots dx_k$$

$$+ \int_{u_1}^{\infty} \int_{u_2}^{\infty} \cdots \int_{u_k}^{\infty} y\, f(x_1, x_2, \ldots, x_k)\, dx_1\, dx_2 \ldots dx_k$$

$$= \pi.$$

The estimate of Π is the random variable $p = \dfrac{\sum\limits_{i=1}^{n} y_i}{n}$, where y_1, y_2, ..., y_n represent a sample of independent random variables defined as in (1) from n vector observations X_1, X_2, ..., X_n randomly sampled from f(X).

If y_1, y_2, ..., y_n are a given set of observations,

$$p = \frac{\sum\limits_{i=1}^{n} y_i}{n} \; .$$

The variance of the binomial variate y is $\sigma^2(y) = \dfrac{\Pi(1 - \Pi)}{n}$.

Its estimate is $\hat{\sigma}^2(y) = \dfrac{p(1 - p)}{n}$,

or, if written in terms of the observations, $s^2(y) = \dfrac{p(1 - p)}{n}$.

To obtain the observed values y_1, y_2, ..., y_n, first, n vectors Z_i of independent random normal deviates are generated; Z has mean vector $m_z = \underline{0}$ and the identity correlation matrix $I = R = Var(Z'Z)$ when $n \to \infty$. Then $X_i = Z_i R^{\frac{1}{2}}$ for i = 1, 2, ..., n, where $R^{\frac{1}{2}}$ is the square root matrix of R. Furthermore, $m_x = \underline{0}$ and $Var(X) = R^{\frac{1}{2}}Z'ZR^{\frac{1}{2}} = R^{\frac{1}{2}}IR^{\frac{1}{2}} = R$ when $n \to \infty$, as is required for the parameters which characterize f(X). If, for each element x_{ij} of any vector X_i, $\ell_j < x_{ij} < u_j$, then y_i is set to 1; but if any element of X_i falls outside the specified interval, y_i is set to 0.

Importance Sampling

An alternate Monte-Carlo approach to integration is referred to as importance sampling. Random numbers upon which the estimates of the unknown parameter are based are generated from a distribution other than the one suggested by the problem. Each value of this "biased" sample is multiplied by an appropriate weighting factor which corrects for having

- 3 -

used the wrong distribution. The probability that a sample will be drawn from an "interesting" or "important" region is increased as a result of the biasing; the probability that it will come from an "uninteresting" or "unimportant" region is decreased--a desirable result. By basing more computations on random numbers generated from portions of the probability region with which the manpower problem is concerned, it is possible to improve the accuracy--that is, reduce the variance--of the estimate.

The model is developed as follows:

Let $g(X) = g(x_1, x_2, \ldots, x_k)$ represent a k-variate uniform probability density function defined over the probability region $A = [(l_1, u_1), (l_2, u_2), \ldots, (l_k, u_k)]$. That is,

$$g(X) \begin{cases} > 0 \text{ if } X \in A \\ = 0 \text{ otherwise} \end{cases}$$

and

$$\int_A g(X)dX = \int_{l_1}^{u_1} \int_{l_2}^{u_2} \cdots \int_{l_k}^{u_k} g(x_1, x_2, \ldots, x_k) \tag{1}$$

Let h represent another function of the random variable X. The expected value of h is defined as

$$E[h(X)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, x_2, \ldots, x_k) \, g(x_1, x_2, \ldots, x_k)dx_1 \, dx_2 \cdots dx_k.$$

In particular, suppose

$$h(X) = \frac{f(X)}{g(X)}$$

with $f(X)$ representing the multivariate normal probability density function whose integral over A is required. Then

$$E\left[\frac{f(X_i)}{g(X_i)}\right] = \int_A \frac{f(X)}{g(X)} \, g(X)dX = \int_A f(X)dX$$

$$= \Pi \tag{2}$$

which is the unknown probability region of interest.

An estimate of Π is the random variable

$$P = \frac{1}{n} \sum_{i=1}^{n} \frac{f(X_i)}{g(X_i)} \tag{3}$$

where $X_1, X_2, \ldots, X_n$ are random vectors sampled from $g(X)$.

- 4 -

Let $y_1, y_2, \ldots, y_n$ where $y_i = \dfrac{f(X_i)}{g(X_i)}$ for any i, represent a sample of independent random variables over A. Then the variance of the random variable p is

$$\sigma^2 (p) = \sigma^2 \left[\frac{1}{n} \sum_{i=1}^{n} \frac{f(X_i)}{g(X_i)} \right] = \frac{1}{n} \sigma^2 \left[\frac{f(X_i)}{g(X_i)} \right] = \frac{1}{n} \sigma^2 [y] . \qquad (4)$$

The variance of y is $\sigma^2 (y) = E(y^2) - E^2(y) = \left[\int_A \frac{f^2(X)}{g^2(X)} g(X) dX \right] - \pi^2 .$

An estimate of $E(y)$ is $\dfrac{1}{n} \sum_{i=1}^{n} \dfrac{f^2(X_i)}{g^2(X_i)}$

where $X_1, X_2, \ldots, X_n$ represent the same sample of random variables uniformly distributed over A and used to estimate π.

Standard Importance Sampling Approach to Integration

Different Monte-Carlo methods for estimating π result from alternate choices for the sampling distribution $g(X)$. In one of the simplest and most frequently used methods of Monte-Carlo sampling, sampling is based on the uniform distribution. Random vectors with limits on the elements defined by the probability region of interest are generated. These observations are then used to compute the ordinate of the given multivariate normal distribution, and a weighted mean of the ordinate values provides the estimate of π. Since random uniform numbers within specified limits can be easily generated, no computational time is wasted through rejection of any sample values. In more detail, let

$$g(X) = g(X_1, X_2, \ldots, X_k) = \begin{cases} \dfrac{1}{\prod\limits_{j=1}^{k} (u_j - l_j)} = \dfrac{1}{[A]} & \text{if } l_j < x_j < u_j \\ 0 \text{ otherwise} \end{cases}$$

This is the k-variate uniform probability density function over the unknown probability region A = $[l_1, u_1), (l_2, u_2), \ldots, (l_k, u_k)]$ with a k-dimensional volumn, the cartesian product of one dimensional intervals (l_j, u_j), j = 1, 2, $\ldots$, k, being denoted by [A]. As is required, $g(X) > 0$ for all X and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(X) dX = \frac{[A]}{[A]} - 1$$

Written in the form of this distribution, the estimate of π is

$$p = \frac{[A]}{n} \sum_{i=1}^{n} f(X_i), \text{ and the variance is } \sigma^2(\nu) + \left[[A] \int_A f^2(X) dX\right] - \pi^2,$$

where X_1, X_2, ..., X_n are random vectors sampled from $g(X)$.

Observations $X_i = [x_{1i}, x_{2i}, ..., x_{ki}]$ which have the $g(X)$ distribution are constructed by first generating n vectors $Z_i = [z_{1i}, z_{2i}, ..., z_{ki}]$ of independent random numbers which are uniform over the interval $(0, 1)$. Then $x_{ji} = (z_{ji})(a_{j, 1_j + 1} - a_{j, 1_j}) + a_{j, 1_j}$ for all i, j. The observed probability written in terms of these observations is

$$p = \frac{|R_f|^{\frac{1}{2}}}{(2\pi)^{k/2}} \frac{\prod_{j=1}^{k}(1_j - u_j)}{n} \sum_{i=1}^{n} \exp\left[-\frac{1}{2}(X_i R_f^{-1} X_i')\right]$$

where R_f is the correlation matrix for the given multivariate normal distribution and $|R_f|$ is the determinant of R.

Evaluating the relative precision of different procedures for estimating π is, of course, equivalent to examining the size of the variances of the different estimates. One way to reduce the variability of the estimate for a given method of Monte-Carlo sampling is to adopt some appropriately designed procedure for stratified sampling. The probability region of interest A is divided into mutually exclusive subregions A_1, A_2, ..., A_t, an independent probability estimate p_s obtained for each A_s, and the sum of these estimates used to estimate the probability over the entire region. If the number of points sampled within each area is proportionate to the size of the region over which sampling is defined, then it can be shown that the variance of the estimate of π for the stratified sample will be less than or equal to, but never greater than, the variance for the estimate obtained without stratification.

There are, of course, alternative approaches to the way in which the total area A is divided into strata. One procedure, particularly appropriate when sampling proceeds over uniform dimensions, is "A Modified Monte-Carlo Quadrature" presented by Haber (2). The area for each stratum is the accumulated product of one-dimensional uniform univariate distributions over k dimensions; i.e.,

$$A_s = \prod_{j=1}^{k} (a_{j, p_j + 1} - a_{j, p_j}) \text{ for } p_j = 0, 1, ..., q_j - 1 \text{ and } s = 1, 2, ..., t.$$

The intervals within each of the k dimensions are constructed so that the differences $a_{j, p_j + 1} - a_{j, p_j}$, are commensurable; in the simplest case,

- 6 -

all of them may be multiples of the smallest one. Consequently, the k-dimensional volumns A_1, A_2, ..., A_t will be commensurable, and the variance of the estimate based on the sum of estimates over t strata will be less than or equal to the variance of an estimate based on an undivided sample; i.e.,

$$\text{Var} \left(\sum_{s=1}^{t} p_s \right) \leq \text{Var} (p).$$

Preliminary experimental results, including results with the advantage of sample stratification, indicated that the uniform distribution was not a suitable sampling distribution for integrating the multivariate normal distribution. The choice of an appropriate design for importance sampling is, of course, not routine, and the variance of the sample estimates may be increased, even infinitely, if the random data points are generated from the wrong distribution. It did appear more profitable to examine in detail experimental results based on a different importance sampling design which showed more promise of having the desired property of reduced sample variance.

Importance Sampling: Boldt Method (Method II)

An original approach to importance sampling was devised by Boldt (3). Boldt suggested that a multivariate normal distribution with a single common factor covariance structure often is a good approximation to covariance structure observed in practice and might serve as a suitable solution for the integral of the sampling distribution over the same limits required for the unknown integral. A simple quadrature solution does exist for variance-covariance matrices with the common factor structure and has been described a number of times in the literature, for example, by Curnow and Dunnett (4) and by Lord (5). Furthermore, the generation of random observations which have one common factor is also quite simple. To force the points within the specified limits of integration complicates the computational procedure somewhat, but does not result in a significant increase in computer time.

Development of the model follows:

Identical with formula (2), the integral over A for the multivariate normal distribution is

$$\pi = \int_A \frac{f(X)}{g(X)} \; g(X)dX = E\left[\frac{f(X)}{g(X)} \right] , \quad \text{where}$$

$\int_A g(X)dX = 1$, and the estimate of π (formula 3) is $p = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \dfrac{f(X_i)}{g(X_i)}$

where X_1, X_2, ..., X_n are random vectors sampled from $g(X)$. Rather than representing $g(X)$ by the uniform distribution, however, sampling is from a rank-one multivariate normal distribution.

- 7 -

Retaining the convention that integration will be performed with respect to standardized variates, let R_g represent a correlation matrix with the structure $r_{ij} = \alpha_i \alpha_j$ $(i \neq j)$, where $-1 \leq \alpha_i \leq +1$ for $i = 1, 2, \ldots, k$. Then

$$
g(X) = \begin{cases} \dfrac{k_g \, |R_g|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(X \, R_g^{-1} \, X')\right]}{|2\pi|^{k/2}} = k_g \, g^*(X) = \dfrac{1}{p_g} \, g^*(X) \text{ if } X \in A; \\[2ex] 0 \text{ otherwise} \end{cases}
$$

The cumulative density function of $g(X)$ can be expressed as a single integral having a product of univariate normal cumulative density functions in the integrand; consequently, p_g is easy to evaluate numerically. In more detail,

$$
\pi = \int_A p_g \, \frac{|R_f|^{\frac{1}{2}}}{|R_g|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}X(R_g^{-1} - R_f^{-1})X'\right] g(X)\, dX \quad \text{and}
$$

$$
p = p_g \frac{|R_f|^{\frac{1}{2}}}{|R_g|^{\frac{1}{2}}} \sum_{i=1}^{n} \exp\left[-\frac{1}{2}X_i(R_g^{-1} - R_f^{-1})X'\right]
$$

where $X_1, X_2, \ldots, X_n$ represent the random sample of vectors from the distribution $g(X)$ defined over A. Based on (4), an estimate of this variance of p is

$$
\hat{\sigma}^2(p) = \frac{1}{n}\left[\frac{1}{n} \sum_{i=1}^{n} \frac{f^2(X_i)}{g^2(X_i)} - \left[\frac{1}{n} \sum_{i=1}^{n} \frac{f(X_i)}{g(X_i)}\right]^2\right]
$$

$$
= \frac{1}{n}\left[\left(\frac{p_g^2}{n} \frac{|R_f|}{|R_g|} \sum_{i=1}^{n} \exp\left[-\frac{1}{2}X_i(R_g^{-1} - R_f^{-1})X'\right]\right) - p^2\right]
$$

where $X_1, X_2, \ldots, X_n$ are the same random vectors sampled from $g(X)$ to estimate π.

- 8 -

<u>Solution for a factored integral</u>. To obtain the estimate for Π as stated in (5), p_g, a solution for the integral of the sampling distribution $g(X)$ over A, the probability region of interest, is required. The sampling distribution $g(X)$ is chosen to have a single common factor correlation structure and, consequently, can be factored into a product of univariate integrals.

In more detail, variates z_1, z_2, ..., z_k which have a multivariate normal distribution with zero means and correlation coefficients $r_{ij} = \alpha_i \alpha_j$ $(i \neq j)$, where $-1 \leq \alpha_i \leq +1$ for $i = 1, 2, \ldots, k$, are generated by the formula $z_j = (1 - \alpha^2_{\;j})^{\frac{1}{2}} x_j + \alpha_j y$. The x_1, x_2, ..., x_k, y are $k + 1$ independent standard normal variates. The cumulative density function of interest for the k variates z_j is

$$F_k(h_{ji}; \{r_{ij}\}) = \text{Prob} \{z_j < h_j \; ; \text{ all } j\}$$

$$= \int_{-\infty}^{h_1} \int_{-\infty}^{h_2} \ldots \int_{-\infty}^{h_j} f(x_1, x_2, \ldots, x_k) dx_1 \, dx_2 \ldots dx_k .$$

Since the x's are mutually independent, F_k can be expressed as the product of single integrals; i.e.,

$$F_k = \int_{-\infty}^{\infty} \text{Prob} \{x_j < (h_j - \alpha_j y) / (1 - \alpha_j^2)^{\frac{1}{2}}; \text{ all } j\} f(y) dy$$

$$= \int_{-\infty}^{\infty} \left[\prod_{j=1}^{k} ((h_j - \alpha_j y) / (1 - \alpha_j^2)^{\frac{1}{2}}) \right] f(y) dy$$

where $f(t) = \exp(-\frac{1}{2} t^2) / (2\pi)^{\frac{1}{2}}$
and $F(t) = \int_{-\infty}^{t} f(t) dt$

This integral is very easy to evaluate numerically by use of some quadrature procedure in which the integral is replaced by a sum. In the present application, the infinite interval $<-\infty, +\infty>$ is taken as $[-5, +5]$ and is divided into an even number of $2m = 100$ intervals with end-points defined by y_0, y_1, ..., y_{2m}. Then, using the sum given by Simpson's rule,

$$F_k (h_{ji}; \{r_{ij}\}) =$$

$$\frac{10}{2m} (\frac{1}{\sqrt{2\pi}}) \{f(y_0) + f(y_{2m}) + 4\left[f(y_1) + f(y_3) + \ldots + f(y_{m-1}) \right]$$

$$+ 2\left[f(y_2) + f(y_4) + \ldots + f(y_{m-2}) \right]\}$$

- 9 -

where $f(y_i) = \exp\left[-\tfrac{1}{2}y^2_i\right] \prod\limits_{j=1}^{k} \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_j} \exp\left[-\tfrac{1}{2} z^2_j\right] dz$

and $z_j = \dfrac{(hj - \alpha_j y_i)}{(1 - \alpha^2_j)^{\frac{1}{2}}}$

for $i = 0, 1, 2, \ldots, 2m$.

The univariate standard cumulative density functions are evaluated by a Hasting's approximation (Hastings, 6).

Since, in most examples, it is not the cumulative density function which must be evaluated, but the density function integrated between any limits ℓ_j and u_j, $j = 1, 2, \ldots, k$, in the infinite range. Therefore, computation of p_g is facilitated by the formula

Prob $\left[\ell_1 \leq x_1 \leq u_1 , \ell_2 \leq x_2 \leq u_2 , \ldots, \ell_k \leq x_k \leq u_k\right]$

$= \Sigma\ (-1)^{\text{no. of lower bounds}}$ Prob $\left[x_1 \leq \Delta_1 , x_2 \leq \Delta_2 , \ldots, x_k \leq \Delta_k\right]$,

all 2^k

possible

combinations

where Δ_j takes on the value l_j or u_j.

Generation of Sample Values. Generation of the Z vectors which have the rank-one multivariate normal distribution and which also lie within the limits defined by A is not a straightforward computational procedure. For the k elements of Z generated from the formula

$z_j = \beta_j x_j - \alpha_j y,$

letting

$\beta_j = (1 - \alpha^2_j)^{\frac{1}{2}},$

it is required that

$l_j \leq z_j \leq u_j,$ where l_j and u_j are the lower and upper bounds defining the $j = 1, 2, \ldots, k$ dimensions for a given probability region A. If $x_1, x_2, \ldots, x_k, y$ are random normal deviates sampled over the infinite domain, the probability that the resulting $z_1, z_2, \ldots, z_k$ will lie within the required region A is Π, and p, an estimate of Π, is determined from the proportion of the sampled vectors which lie within the region A. It is desirable, however, to generate Z variates which have the rank-one

- 10 -

distribution without rejecting any observations. If, for instance, A is
very small, the number of observations which are rejected will be very
large; the associated amount of computation will also be very large.
Alternatively, Z values can be constrained to lie within the specified
limits by appropriately scaling the range of the variates
$x_1, x_2, \ldots, x_k, y$. Since

$$\pi = \text{Prob}\left[\{\ell_j \leq \beta_j x_j + \alpha_j y \leq u_j\}; \text{ all } j\right]$$

$$= \text{Prob}\left[\left\{\frac{\ell_j - \alpha_j y}{\beta_j} \leq x_j \leq \frac{u_j - \alpha_j y}{\beta_j}\right\}; \text{ all } j\right]$$

$$= \text{Prob}\left[\{\ell_j^* \leq x_j \leq u_j^*\}; \text{ all } j\right], \quad \text{letting} \quad \ell_j^* = \frac{\ell_j - \alpha_j y}{\beta_j} \text{ and } u_j^* = \frac{u_j - \alpha_j y}{\beta_j}, \quad \cdot$$

the variables $x_1, x_2, \ldots, x_k$ can be constrained to lie within the required
range by first sampling the random variable y anywhere over the infinite
domain and then constructing limits for the x's as a function of y. This,
of course, is saying that

$$\text{Prob}\left[\{\ell_j^* \leq x_j^* \leq u_j \mid y\}; \text{ all } j\right] = 1,$$

which is not the required π, nor does such a method yield the correct
distribution of points for the sampling probability density function g.

To clarify the discussion to follow, suppose y can take on exactly
t possible value s, each with probability π_t. Then the unknown

$$\pi = \sum_{s=1}^{t} \pi_s.$$

To generate points having the correct distribution, if n represents a
very large number of points,

$$n \cdot \frac{\pi_s}{\sum_{s=1}^{t} \pi_s} = n_s \text{ points}$$

must be generated for each value of s. It is convenient to represent
each π_s in terms of conditional probabilities:

$$\pi = \text{Prob}\left[\{\ell_j \leq z_j \leq u_j \mid y = s\}; \text{ all } j\right]. \quad \text{Prob}\left[y = s\right]$$

$$= q_s \cdot r_s .$$

- 11 -

The conditional probability term q_s can be easily computed:

$$q = \text{Prob}\left[\{\ell_j \le z_j \le u_j| y = s\}; \text{ all } j\right]$$

$$= \text{Prob}\left[\{\frac{\ell_j - \alpha_j y}{\beta_j} \le x_j \le \frac{u_j - \alpha_j y}{\beta_j}| y = s\}; \text{ all } j\right]$$

$$= \left[\frac{\ell_j - \alpha_j y}{\beta_j} \le x_j \le \frac{u_j - \alpha_j y}{\beta_j} | y = s\right]$$

since $x_1, x_2, \ldots, x_k$ are independent standard normal variates and q_t is simply the accumulated product of probabilities associated with bounds of normal deviates for independent standard normal univariate distributions. Furthermore, assume that all

Prob $[y = s]$ = Prob $[y = s']$ for s, s' = 1, 2, ..., t; i.e., that all values of y are equally likely. Then

$$\pi_s = \frac{q_s \cdot r_s}{\displaystyle\sum_{s=1}^{t} q_s \cdot r_s} = \frac{q_s}{\displaystyle\sum_{s=1}^{t} q_s}$$

and

$$n_s = \frac{n \cdot q_s}{\displaystyle\sum_{s=1}^{t} q_s} .$$

Aside from purposes of explanation, the variate y is, of course, continuous. Rather than t discrete values which y can take, there are t intervals or ranges of values, each bounded by a lower and upper limit v_s and v_{s+1}, say. Then

$$\text{Prob}\left[\{\ell_j \le s_j \le u_j| v_s \le y \le v_{s+1}\}; \text{ all } j\right]$$

represents a range of probabilities corresponding to the range of y. To estimate this range, an approximation must be introduced, letting one probability q_s, say, represent the range of possible values. If sufficiently small intervals are constructed for y, the probabilities will have a small range for each interval, and the mean of the smallest and largest probabilities should yield a good approximation; alternatively, y could be chosen as the normal deviate corresponding to the probability midpoint of each interval. The latter alternative was chosen for the experiments described here.

THE MONTE-CARLO EXPERIMENTS

Procedure

The relative efficiencies of the different Monte-Carlo methods were examined for a variety of experimental problems. Two sets of runs were prepared, one with four variates and the other with eight; the experimental problems for the two sets were designed to be somewhat comparable. The correlation matrix R_f characterizing the multivariate normal distribution for the "unknown" integral was chosen to have a single common factor with all correlations of the structure $r_{ij} = \alpha_i \alpha_j$, where

$-1 \leq \alpha_i \leq +1$ for $i = 1, 2, \ldots, k$. Then the unknown probability area could be verified from the same quadrature formula used earlier to compute the probability area for the single common factor distribution from which the random observations are generated. Such a choice for R_f, however, lessens the representiveness of the problem. To evaluate the importance sampling approach, results should be compared when sampling proceeds from a distribution which poorly approximates the distribution of interest as well as when both distributions each have only one common factor. To increase the generality of the results, therefore, for each single common factor R_f, both a "good" single factor approximation and a "poor" single factor approximation were constructed for the sampling distribution.

All correlations in R_f were chosen to be .90 (both for four and eight variates) for one set of problems. Such high positive correlations will yield multidimensional probability regions highly concentrated about the center of the distribution. To work with probability regions more evenly distributed over the infinite domain, all correlation elements were set to .10. Four- and eight-variate matrices containing randomly selected positive correlations with more than one common factor were also constructed, although a quadrature procedure for checking the Monte-Carlo estimates was not available. For the two kinds of single common factor correlation matrices used for sampling, one R_g had the structure $r_{ij} = \alpha_i \alpha_j = (.90)(.90) = .81$, and the other had the structure $r_{ij} = (.10)(.10) = .01$ (both for four- and eight-variate problems). For the "unknown" integral with all $r_{ij} = .90$, the R_g with all correlations equal to .81 represented the "good" importance sampling approximation; the R_g with all correlations equal to .01 represented the "poor" approximation. Conversely, when all r_{ij} were set to .10 in R_f, the R_g with correlations equal to .81 was the poor sampling approximation and the R_g with correlations of .01 was the good one.

- 13 -

Probability estimates were also obtained for 3 different kinds of integration limits. One set of limits included only the central portion of the distribution, -1. to +1. for all four or eight variates. Another set included the tails, $-\infty$ to 0. for all variates. In a third set, the limits contained a variety of both tails and central portions (-5. to -1.; -1. to 0; 0 to 1; and 1. to 5. for four variates and -3. to 1; -3 to 2; -3 to 3.; -2 to 1.; -2 to 2.; -2 to 3.; -1. to 2.; -1 to 3. for eight variates).

In summary, the series of experiments used to compare the relative efficiency of two Monte-Carlo methods for evaluating the multivariate normal integral contained four categories of independent variable: 1) number of variates, 4 and 8; 2) structure of the multivariate normal distribution of interest, as characterized by a one common-factor correlation matrix with high positive correlations, a one common-factor correlation matrix with low positive correlations, and a correlation matrix with randomly selected positive correlations; 3) goodness of approximation of the sampling distribution to the multivariate normal distribution being integrated; 4) range of integration, with limits involving only the central portion of the distribution or including both tail and central portions as well.

All possible combinations of the independent variables planned for the series of experiments totaled 36 problems; 18 of these problems were based on four-variate distributions and 18 comparable problems were for the eight-variate distributions. For each of these problems, 10 independent probability estimates were obtained. The estimates for the four-variate problems were each based on an n of 1000 random vector observations; 10,000 random vectors were generated for each of the eight-variate problems.

Results from the Monte-Carlo Experiments

Results are presented in Tables 1, 2, and 3. For each of the four- and eight-variate series, 9 different probability regions were being evaluated. The first line of values presented for each such region is based on the first and simplest method of Monte-Carlo sampling described (Method I). The second two lines of values were obtained using the Boldt importance sampling approach (Method II), but with one set of estimates being obtained from a "good" sampling distribution approximation and the other set from a "poor" approximation.

For the single common-factor distribution, probabilities computed by quadrature are presented in the tables as Π, to represent the "population" value. The average squared deviation from this population value $(s^2|\Pi)$ over the 10 estimates per problem was used as a measure of the accuracy of a given method[1]. The average squared deviation from the observed mean

---

[1] It is the within sample variances which are presented in formulas (4), (8), and (9). The measures of between sample variability, however, were more discriminatory of the effectiveness of the different Monte-Carlo methods and are therefore the ones presented in the tables.

$(s^2|\bar{p})$ based on each set of 10 samples was also computed. It is the square roots of these measures ($s|\pi$ and $s|p$, "standard deviations") that are presented in Tables 1 and 2. The standard deviations obtained for the simple Monte-Carlo method were used as the baselines by which relative amount of variation for different problems could be evaluated. These ratios of standard deviations are also presented in Tables 1 and 2 for each set of problems.

The first observation which should be made about the results is an apparent equivalence of the two measures of variability, $s|\pi$ and $s|\bar{p}$. That is, the degree to which one Monte-Carlo method is more precise than another for a given problem appears to be independent of whether the deviations of the estimates are taken about the observed mean or about the population value. Of course, such a generalization can be made only with respect to unknown integrals based on single common factor distributions. Equivalence of the measures, if equivalent over all types of problem, could be taken as an indication that the Monte-Carlo estimates, even though highly variable, are unbiased estimates and can be expected under increased sampling to converge to the true population value.

To further clarify the form of the results, the ratios of standard deviations were extracted from Tables 1 and 2 and regrouped to form Table 3. Noting that a ratio of standard deviations above 1.0 indicates superiority of the importance sampling method, whereas a ratio below 1.0 favors the simpler sampling procedure, a striking characteristic of Table 3 is that neither method is consistently superior over the different types of integral being evaluated. The ideal result, of course, is to find some method which yields estimates with a marked reduction of variance on all problems. With the results shown here, the types of problem for which one method might be superior to another is a matter only for hypothesizing. Ultimately, perhaps, the most economical procedure with respect to computer time will be the design of some test in which the type of probability integral to be evaluated is examined by the computer before it proceeds to analysis by one of several alternate methods. At most, these experimental results may have bearing on some entirely new approach to the problem of evaluating multivariate normal integrals.

An independent research project is being conducted by Cecil Johnson (7) at BESRL on increasing the goodness of fit of the single common-factor sampling distribution to the multivariate normal distribution of interest to increase the precision of importance sampling. Preliminary results of this research indicate that using improved methods to determine the parameters from which the common factor random entities are generated can result in an appreciable reduction in variance. That such an approach is a fruitful one is supported by the data presented in the present publication. In fact, the one striking observation which comes from Tables 1, 2, and 3 is the superiority of the "good" approximation sampling distribution to the "poor" approximations in yielding minimum variance estimates. This advantage shows up in all 24 of the problems where goodness of fit was compared and goes as high as 42 to 1 in problem 22

- 15 -

with eight variates. One should be aware, of course, of the fact that, when the sampling approximation is very good (i.e., when the sampling distribution is nearly identical to the distribution associated with the unknown integral), the variable components on which an importance sampling estimate is based are very small relative to the scaling constant computed by quadrature. Therefore, the superiority shown in problems 1, 4, 7, 10, 13, 16, 22, 28, and 34 is partly a consequence of the fact that the unknown distribution is one of the rare distributions whose integral value is near to the value computed by the given quadrature formula, a situation not to be encountered often. The real test of an importance sampling approach is whether an advantage can be observed when the sampling approximation is only moderately good or poor. Such an advantage does show up in problems 3 and 21 (for four and eight variates) and 9 (for four variates), both poor approximation problems, but not in other poor approximation problems, 2, 8, 15, 20, 26, and 33, (for four and eight variates) and 27 and 32 (for eight variates).

The determination of a suitable single common-factor distribution and the manner in which random entities are generated from this distribution is not a straight-forward procedure. By varying such a procedure, precision of the estimates based on even a poor sampling distribution can be improved. The results presented here may be considered relatively crude with respect to this aspect of the problem; improvement is expected when the procedures under investigation by Johnson are incorporated into the techniques described here.

The estimates in Table 3 do appear to favor slightly the importance sampling approach when the region of integration is over the center of the distribution. The best results, in problems 4 and 22 for four and eight variates, were obtained when the range of integration was from -1. to +1., the center, even though the associated multivariate normal distribution was least concentrated about the center (i.e., all r's = .10). Furthermore, the ratios of standard deviations were greater than one in all but one of the problems designed for the central region, regardless of the type of distribution being integrated. The importance sampling method tends to be relatively less efficient when tails are included in the integration limits. For example, problems 23 and 24 for eight variates indicate a gross failure of the imporatnce sampling method, although these problems are also based on poor approximations to eight variate common factor distributions. Furthermore, any observation on tail results cannot be stated too conclusively, since the tail regions for the data under discussion were not examined to the exclusion of any of the more central regions.

Clearly, the relative advantage of one sampling method over another is highly sensitive to variations in the region of integration relative to the structure of the correlation matrix. For instance, the ratios for the four-variate distribution characterized by a single common-factor correlation matrix with large positive elements (problems 1, 7, 13, 19, 25, and 31) increased by a multiple of about 2 when the integration limits included a variety of ranges, even though the probability area was small. By contrast, when the correlation matrix had very small positive elements, the advantage of the importance sampling method decreased to about one-half when the tails were included (problems 4, 10, 16, 22, 28, and 34).

- 16 -

Table 1

MONTE-CARLO ESTIMATES BASED ON TWO DIFFERENT SAMPLING PROCEDURES FOR EVALUATING THE MULTIVARIATE NORMAL INTEGRAL

FOUR-VARIATE PROBLEMS
$(N = 1000)$

	All correlations = .90; Centered limits: -1 to +1						
Problem	Method	π	p	$s \mid p$	s ratio	$s \mid \pi$	s ratio
1	I	.5114	.5121	.02175		.02177	
1	II $r_g=.81$ Good	.5114	.5125	.00871	2.50	.00877	2.48
2	II $r_g=.01$ Poor	.5114	.5125	.02836	0.77	.02838	0.77

	All correlations = .10; Centered limits: -1 to +1						
Problem	Method	π	p	$s \mid p$	s ratio	$s \mid \pi$	s ratio
3	I	.2202	.2203	.00331		.01331	
3	II $r_g=.81$ Poor	.2202	.2156	.01092	1.22	.01 36	1.12
4	II $r_g=.01$ Good	.2202	.2204	.00037	35.56	.00040	33.46

	Full Rank, positive correlations; Centered limits: -1 to +1			
Problem	Method	p	$s \mid p$	s ratio
5	I	.3217	.00212	
5	II $r_g=.81$	.3187	.01645	1.29
6	II $r_g=.01$	.3226	.00323	6.56

	All correlations = .90; limits include tails: $-\infty$ to 0						
Problem	Method	π	p	$s \mid p$	s ratio	$s \mid \pi$	s ratio
7	I	.3693	.3618	.01639		.01803	
7	II $r_g=.81$ Good	.3693	.3705	.00688	2.38	.00698	2.58
8	II $r_g=.01$ Poor	.3693	.3296	.03068	.53	.05018	.36

Table 1 (continued)

All correlations = .10; limits include tails: -∞ to 0

Problem	Method	π	p	s$\mid$p	s ratio	s$\mid\pi$	s ratio
9	I	.0871	.0869	.01065		.01066	
9	II r_g=.81 Poor	.0871	.0843	.00963	1.11	.01003	1.03
10	II r_g=.01 Good	.0871	.0871	.00075	14.23	.00075	14.22

Full rank, positive correlations; limits include tails: -∞ to 0

Problem	Method	p	s$\mid$p	s ratio
11	I	.1951	.01304	
11	II r_g=.81	.1954	.01783	.73
12	II r_g=.01	.1919	.00892	1.46

All correlations = .90; limits are varied

Problem	Method	π	p	s$\mid$p	s ratio	s$\mid\pi$	s ratio
13	I	.07054	.07410	.00780		.00858	
13	II r_g=.81 Good	.07054	.07063	.00145	5.38	.00145	5.90
14	II r_g=.01 Poor	.07054	.06926	.00723	1.08	.00734	1.17

All correlations = .10; limits are varied

Problem	Method	π	p	s$\mid$p	s ratio	s$\mid\pi$	s ratio
15	I	.1535	.1544	.01547		.01550	
15	II r_g=.81 Poor	.1535	.1463	.02461	.63	.02562	.60
16	II r_g=.01 Good	.1535	.1534	.00042	36.74	.00043	35.93

Full rank, positive correlations; limits are varied

Problem	Method	p	s$\mid$p	s ratio
17	I	.1188	.01260	
17	II r_g=.81	.1307	.04198	.30
18	II r_g=.01	.1148	.00320	3.94

Table 2

MONTE-CARLO ESTIMATES BASED ON TWO DIFFERENT SAMPLING PROCEDURES FOR EVALUATING THE MULTIVARIATE NORMAL INTEGRAL

EIGHT-VARIATE PROBLEMS
$(N = 10,000)$

All correlations = .90; centered limits: -1 to +1

Problem	Method	Π	p	s|p	s ratio	s|Π	s ratio
19	I	.4302	.4309	.00335		.00343	
19	II r_g=.81 Good	.4302	.4304	.00327	1.03	.00327	1.05
20	II r_g=.01 Poor	.4302	.4245	.01582	.21	.01682	.20

All correlations = .10; centered limits: -1 to +1

Problem	Method	Π	p	s|p	s ratio	s|Π	s ratio
21	I	.0499	.0502	.00229		.00232	
21	II r_g=.81 Poor	.0499	.0502	.00214	1.07	.00216	1.08
22	II r_g=.01 Good	.0499	.0499	.00005	45.02	.00006	38.87

Full rank, positive correlations; centered limits: -1 to +1

Problem	Method	p	s|p	s ratio
23	I	.1757	.00378	
23	II r_g=.81	.1731	.00291	1.30
24	II r_g=.01	.1751	.00238	1.59

All correlations = .90; limits include tails: $-\infty$ to 0

Problem	Method	Π	p	s|p	s ratio	s|Π	s ratio
25	I	.3211	.3206	.00412		.00415	
25	II r_g=.81 Good	.3211	.3212	.00280	1.47	.00280	1.48
26	II r_g=.01 Poor	.3211	.2590	.02346	.17	.06640	.06

Table 2 (continued)

All correlations = .10; limits include tails: $-\infty$ to 0

Problem	Method	Π	p	s|p	s ratio	s|Π	s ratio
27	I	.0141	.0141	.00095		.00095	
27	II r_g=.81 Poor	.0141	.0126	.00136	.70	.00203	.47
28	II r_g=.01 Good	.0141	.0141	.00010	9.03	.00011	8.64

Full rank, positive correlations; limits include tails: $-\infty$ to 0

Problem	Method		p	s|p	s ratio
29	I		.1349	.00337	
29	II r_g=.81		.1249	.00773	.44
30	II r_g=.01		.1206	.00697	.48

All correlations = .90: limits are varied

Problem	Method	Π	p	s|p	s ratio	s|Π	s ratio
31	I	.5962	.5960	.00564		.00564	
31	II r_g=.81	.5962	.5966	.00552	1.02	.00553	1.02
32	II r_g=.01	.5962	.5895	.08477	.07	.08503	.07

All correlations = .10; limits are varied

Problem	Method	Π	p	s|p	s ratio	s|Π	s ratio
33	I	.4359	.4393	.00379		.00510	
33	II r_g=.81	.4359	.3128	.03915	.10	.12919	.04
34	II r_g=.01	.4359	.4363	.00144	2.64	.00148	3.45

Full rank, positive correlations: limits are varied

Problem	Method		p	s|p	s ratio
35	I		.5234	.00488	
35	II r_g=.81		.4900	.09384	.05
36	II r_g=.01		.5191	.02511	.19

Table 3

STANDARD DEVIATION RATIOS: VARIABILITY OF IMPORTANCE SAMPLING PROBABILITY ESTIMATES RELATIVE TO METHOD I MONTE-CARLO ESTIMATES

Correlation Structure		Varied Integration Limits					
$\{r_f\}$	$\{r_g\}$	Problem No.	S.D. Ratio	Problem No.	S.D. Ratio	Problem No.	S.D. Ratio
		Four Variates (N = 1000)					
.90	.81, good	1	2.50	7	2.38	13	5.38
.90	.01, good	2	.77	8	.53	14	1.08
.10	.81, poor	3	1.22	9	1.11	15	.63
.10	.01, good	4	35.56	10	14.23	16	36.74
Random	.81	5	1.27	11	.73	17	.30
Random	.01	6	6.56	12	1.46	18	3.94
		Eight Variates (N = 10,000)					
.90	.81, good	19	1.03	25	1.47	31	1.02
.90	.01, poor	20	.21	26	.17	32	.07
.10	.81, poor	21	1.07	27	.70	33	.10
.10	.01, good	22	45.02	28	9.03	34	2.64
Random	.81	23	1.30	29	.44	35	.05
Random	.01	24	1.59	30	.48	36	.19

Second Series of Monte-Carlo Experiments

An additional series of experiments, similar to the previous series, was designed to examine the precision of Monte-Carlo estimates for very small probability regions. A variety of single-factor problems, for which Π could be easily approximated by quadrature, was evaluated on the computer until a set of probability regions similar in magnitude, but small, were found. The largest of these regions was .02965 and the smallest was .00007. One set of experiments was designed for four variates and another for eight variates. The correlation matrices characterizing the distributions to be integrated were the same single common-factor matrices used for the previous problems; all correlation coefficients in these matrices were equal either to .90 or to .10.

In the second experimental series, results were highly encouraging in favor of the importance sampling approach. With small Π, the ratios of standard deviations measuring precision of Monte-Carlo Method II, relative to Method I, were quite large, ranging from 7 to 1 to as high as 2989 to 1. The best result for the problems previously discussed was only 45 to 1 (problem 22, Table 3). Furthermore, the direction of the results was consistent. For every problem, Method II was more precise than Method I. The two variance measures, with deviations taken about the observed mean in one measure and about the population parameter in the other, gave comparable results, as they did in the first series of problems. This conformity was again taken to be an indication of lack of bias in the results.

Judging by an inspection of the results of Tables 4 and 5, the advantage favoring the importance sampling method appears to be related only to the size of the probability region. Central regions versus tail regions, or four versus eight variates, do not appear to distinguish one method from another. In general, better results were obtained with Method II when the sampling distribution approximation was good, although exceptions occurred in problems 42, 44, and 46.

The point where the advantage due to size of Π disappears is, of course, not entirely clear. The very large ratios of Tables 4 and 5 are for Π's of .003 or less. Note that in problem 47, when Π is as large as .030, the ratio drops to only 7 to 1. The ambiguous results presented in Tables 1 and 2, with Method I being superior in some problems and Method II being superior in others, usually involved very large Π's (e.g., .5114 or .4302). But when Π was small (.0499 or .0705), as in problems 13, 14, 21, and 22, results favored the importance sampling approach. On the other hand, Π for problem 27 was small, .0141, and the relative sizes of the variances indicate greater precision in the brute-force method. Very likely, some interaction is occurring between size of the probability region and adequacy of the sampling distribution approximation. Results for Method II tended to be poor when integration was over the tails of a relatively flat multivariate normal distribution (such as when all r's equaled .10), and sampling was from a distribution with most of its mass concentrated over the center (as when all r's = .90); observe problems 9 and 27.

Table 4

COMPARISON OF MONTE-CARLO ESTIMATES FOR SMALL PROBABILITY REGIONS IN EITHER THE TAILS OR THE CENTER OF THE MULTIVARIATE NORMAL DISTRIBUTION

FOUR-VARIATE PROBLEMS
$(N = 1000)$

All correlations = .90; Central limits: 0 to .30

Problem	Method	Π	p	$s\|p$	s ratio	$s\|\Pi$	s ratio
37	I	.00298	.00370	1.7349×10^{-3}		1.8776×10^{-3}	
37	II r_g=.81 Good	.00298	.00296	2.7513×10^{-6}	630.6	1.7728×10^{-5}	105.9
38	II r_g=.01 Poor	.00298	.00296	6.2519×10^{-6}	277.5	1.9058×10^{-5}	98.5

All correlations = .10; Central limits: 0 to .30

Problem	Method	Π	p	$s\|p$	s ratio	$s\|\Pi$	s ratio
39	I	.00020	.00000	0		2.0100×10^{-4}	
39	II r_g=.81 Poor	.00020	.00020	1.8481×10^{-7}	0	1.5733×10^{-6}	127.8
40	II r_g=.01 Good	.00020	.00020	3.8535×10^{-8}	0	1.6129×10^{-6}	124.6

All correlations = .90; Tail limits: 1.9 to 2.5

Problem	Method	Π	p	$s\|p$	s ratio	$s\|\Pi$	s ratio
41	I	.00266	.00100	1.0954×10^{-3}		1.9855×10^{-3}	
41	II r_g=.81 Good	.00266	.00267	8.9168×10^{-6}	122.8	1.3949×10^{-5}	142.3
42	II r_g=.01 Poor	.00266	.00265	5.9575×10^{-5}	18.4	5.9927×10^{-5}	33.1

All correlations = .10; Tail limits: 1.5 to 2.5

Problem	Method	Π	p	$s\|p$	s ratio	$s\|\Pi$	s ratio
43	I	.00007	.00010	3.0000×10^{-4}		3.0170×10^{-4}	
43	II r_g=.81 Poor	.00007	.00007	8.9698×10^{-7}	334.4	1.5096×10^{-6}	199.9
44	II r_g=.01 Good	.00007	.00007	5.7936×10^{-7}	517.8	7.1184×10^{-7}	423.8

Table 5

COMPARISON OF MONTE-CARLO ESTIMATES FOR SMALL PROBABILITY REGIONS IN EITHER
THE TAILS OR THE CENTER OF THE MULTIVARIATE NORMAL DISTRIBUTION

EIGHT-VARIATE PROBLEMS
($N = 10,000$)

All correlations = .90; Central limits: 0 to .35

Problem	Method	Π	p	$s\mid p$	s ratio	$s\mid\Pi$	s ratio
45	I	.00012	.00013	1.1874×10^{-4}		1.1945×10^{-4}	
45	II $r_g=.81$ Good	.00012	.00012	7.5737×10^{-8}	1567.8	8.2030×10^{-7}	145.6
46	II $r_g=.01$ Poor	.00012	.00012	1.4855×10^{-7}	799.3	8.0974×10^{-7}	147.5

All correlations = .10; Central limits: 0 to .85

Problem	Method	Π	p	$s\mid p$	s ratio	$s\mid\Pi$	s ratio
47	I	.02965	.02980	1.6328×10^{-3}		1.6398×10^{-3}	
47	II $r_g=.81$ Good	.02965	.02963	1.0044×10^{-4}	16.3	1.0293×10^{-4}	15.9
48	II $r_g=.01$ Poor	.02965	.02959	2.1859×10^{-4}	7.5	2.2927×10^{-4}	7.2

All correlations = .10; Central limits: 0 to .85

Problem	Method	Π	p	$s\mid p$	s ratio	$s\mid\Pi$	s ratio
49	I	.00010	.00013	9.0000×10^{-5}		9.4868×10^{-5}	
49	II $r_g=.81$ Poor	.00010	.00010	3.6664×10^{-7}	245.5	4.6482×10^{-7}	204.1
50	II $r_g=.01$ Good	.00010	.00010	3.4429×10^{-8}	2614.1	1.9871×10^{-7}	477.4

All correlations = .90; Tail limits; 1.8 to 2.5

Problem	Method	Π	p	$s\mid p$	s ratio	$s\mid\Pi$	s ratio
51	I	.00095	.00088	2.8214×10^{-4}		2.9143×10^{-4}	
51	II $r_g=.81$ Good	.00095	.00095	2.4341×10^{-6}	115.9	2.5709×10^{-6}	113.4
52	II $r_g=.01$ Poor	.00095	.00096	9.7721×10^{-6}	28.9	1.0121×10^{-5}	28.8

All correlations = .10; Tail limits: 0.8 to 2.5

Problem	Method	Π	p	$s\mid p$	s ratio	$s\mid\Pi$	s ratio
53	I	.00008	.00009	8.3066×10^{-5}		8.4077×10^{-5}	
53	II $r_g=.81$ Poor	.00008	.00008	1.5748×10^{-6}	52.7	1.5968×10^{-6}	52.6
54	II $r_g=.01$ Good	.00008	.00008	3.2392×10^{-7}	256.4	3.5284×10^{-7}	238.3

One implication of the results obtained for small Π's is that the recommended importance sampling approach could be improved by introducing sample stratification. The probability region A would be divided into mutually exclusive regions A_1, A_2, ..., A_t, and the final estimate p would be the sum of t independent estimates p_1, p_2, ..., p_t; i.e.,

$$\hat{\Pi} = p = \sum_{s=1}^{t} p_s.$$ Some properties of sample stratification were presented under the section describing importance sampling based on a uniform sampling distribution. The advantage stated there was that, when the number of points sampled in each region is proportional to the magnitude of the region, the within sample variance based on stratification will be less than or equal to the variance for an unstratified sample. The importance sampling method used in the experiments should have an additional advantage because the probability regions will be small.

SUMMARY

Experimental results have been presented to evaluate two different methods of Monte-Carlo sampling to integrate the multivariate normal distribution. Using random vector observations generated to have the distribution of interest, one method is basically a count of the observations which lie within the region of integration. A more complex method is an adaptation of importance sampling in which a single common-factor multivariate normal distribution is the sampling distribution. Random vector entities are constrained in such a way that only observations which lie within the specific integration limits are generated.

The precision of the estimates for the two methods was compared from the magnitude of the variances of the estimates over independent samples. The form of the results appears to be affected both by the size of the probability region over which integration is performed and also by the goodness of fit of the importance sampling distribution to the distribution under evaluation. When the probability region is very small, the importance sampling approach is clearly the more efficient method. For larger probability regions, the importance sampling approach is superior to the tallied method when the sampling approximation is a good approximation. When the importance sampling approximation is poor, the precision of the probability estimate favors the simpler Monte-Carlo procedure. The results indicate that applying the importance sampling technique to a probability region divided into strata, a very small probability being associated with each stratum, will yield precise estimates for a wide variety of problems. Sampling distributions other than the multivariate normal might also be superior for selected portions of the probability region. Finally, BESRL research scientists are continuing studies to improve the adequacy of the approximation of the one common-factor distribution to a multivariate normal distribution of interest. Information gained from the new research may be expected to increase the efficiency of the techniques described here.

LITERATURE CITED

1. Hillier, Ann. A program for computing probabilities over rectangular regions under the multivariate normal distribution. Technical Report No. 54. Applied Mathematics and Statistics Laboratories. Stanford University, California. 1961.

2. Haber, Seymour. A modified Monte-Carlo quadrature. _Mathematics of Computation_. 1966. 20, 361-368.

3. Boldt, Robert. An importance sampling integration of the multivariate normal curve. U. S. Army Behavioral Science Research Laboratory. (In preparation).

4. Curnow, R. N. and C. W. Dunnett. The numerical evaluation of certain multivariate normal integrals. _Annals of Mathematical Statistics_. 1962. 33, 571-579.

5. Lord, Frederic M. Evaluating the multivariate normal integral over a rectangular region when the correlation matrix has only one common factor. _Research Bulletin_. Educational Testing Service. November 1957.

6. Hastings, Cecil, Jr. Approximations for digital computers. Princeton, N. J. 1955.

7. Johnson, Cecil. Programming plan and supplement number 1. OVHD-02, Statistical Research and Analysis Division, U. S. Army Behavioral Science Research Laboratory. 1966.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
U. S. Army Behavioral Science Research Laboratory, Wash., D. C.	Unclassified
	2b. GROUP

3. REPORT TITLE

EXPERIMENTAL COMPARISON OF MONTE-CARLO SAMPLING TECHNIQUES TO EVALUATE THE MULTIVARIATE NORMAL INTEGRAL

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Elizabeth N. Abbe

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1969	38	7

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. DA R&D Proj. No. 2Q062106A723	Technical Research Note 208
c. Optimization Models	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d. b-12	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office, Chief of Research and Development, DA, Washington, D. C.

13. ABSTRACT

The research effort of the OPTIMIZATION MODELS Work Unit is concerned with providing means of solution to personnel management problems relating to distribution, training, career progression, reassignment, and utilization of manpower in current and future Army personnel subsystems. The present study compares techniques for estimating manpower requirements where a number of individually varying skills, performance potentials, background and behavioral factors must be considered. The specific objective was to evaluate two different numerical methods for estimating probability when a multivariate normal model (for example, one involving scores on a battery of tests) can be assumed.

In a series of simulation experiments in which random vector observations were generated, probability estimates were computed by each of the two methods. Probability regions on which the experiments were based were chosen to have a variety of properties. The precision of the two methods was compared from the magnitudes of the variances of the probability estimates over independent samples. Results indicated that when the probability region is very small, the more complex of the two methods (importance sampling) is superior; but when the sampling approximation is poor, the precision of the probability estimates favors the simpler Monte-Carlo procedure. The computational procedures developed appear to be practical methods of estimating probability based on multiple scores for individuals in a sample population.

DD FORM 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

- 29 -

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
*Multivariate analysis						
Statistics						
*Sampling techniques						
*Probability estimates						
*Multivariate model						
Simulation experiments						
Vectors						
*Probability regions						
*Computational procedures						
*Monte-Carlo methods						
Optimization techniques						